UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/997,957	11/30/2001	Joachim Frank	DE920000055US1 (590.080)	5057

35195 7590 06/01/2005

FERENCE & ASSOCIATES
409 BROAD STREET
PITTSBURGH, PA 15143

EXAMINER	
VO, HUYEN X	

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 06/01/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/997,957	FRANK ET AL.	
	Examiner	Art Unit	
	Huyen Vo	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *16 March 2005*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *19-49* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *19-49* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *30 November 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1. Applicant has submitted an amendment filed 3/16/2005, amending claims 19, 25, 31, 39, and 47-49, while arguing to traverse the art rejection based on an amended limitation regarding "*identifying a known speaker from among the plurality of speakers*" (*see claim amendment*). Applicant's arguments have been fully considered but they are not persuasive. Kimber et al. (US 5598507) teach the step of identifying a speaker from among a plurality of speaker by using speech models trained by said plurality of speakers (*col. 7, lines 41-67, recognizing speaker based on "usual speaking style of the speaker—the actual words used are unimportant"*). As thus, previous ground of rejection is maintained.

Claim Rejections - 35 USC § 102

2. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless – (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

3. Claims 25-30, 39-41, 43-47, and 49 are rejected under 35 U.S.C. 102(b) as being anticipated by Kimber et al. (US Patent No. 5598507).

4. Regarding claim 39, Kimber et al. disclose an apparatus for processing a continuous audio stream containing human speech from a plurality of speakers related

to at least one particular transaction, comprising: a predeterminer which predetermines at least one known speaker from among the plurality of speaker (*element 212 in figure 12, initial training of speaker models HMMs and/or col. 7, lines 41-67, recognizing speaker based on "usual speaking style of the speaker—the actual words used are unimportant"*); a detector which detects speaker changes in the audio stream (*col. 11, ln. 38-67*); a recognizer which recognizes the predetermined speaker in the audio stream (*col. 11, ln. 38 to col. 12, ln. 27*); an indexer for indexing the audio stream dependent on a detected speaker change and a recognized predetermined speaker (*fig. 12 or col. 11, ln. 38 to col. 12, ln. 27*).

5. Regarding claims 25, 43, and 49, Kimber et al. disclose a method, apparatus, and program storage device readable by machine for processing a continuous audio stream containing human speech of a plurality of speakers related to at least one particular transaction, comprising the steps of: identifying a known speaker from among the plurality of speakers (*col. 7, lines 41-67, recognizing speaker based on "usual speaking style of the speaker—the actual words used are unimportant"*); digitizing the continuous audio stream (*ADC is inherently included in a computer system of figure 11*); detecting a speaker change in the digitized audio stream (*col. 11, ln. 38-67*); performing a speaker recognition if a speaker change is detected (*col. 11, ln. 38 to col. 12, ln. 27*); indexing the audio stream with respect to the detected speaker change if the known speaker is recognized (*fig. 12 or col. 11, ln. 38 to col. 12, ln. 27*).

6. Regarding claims 26 and 44, Kimber et al. further disclose a method and apparatus according to claims 25 and 39, comprising the further step of protocolling time information for detected speaker changes (*col. 11, ln. 21-37*).

7. Regarding claims 27 and 40, Kimber et al. further disclose a method and apparatus according to claims 25 and 31, wherein the step of detecting a speaker change and/or the step of performing a speaker recognition is/are preceded by the further step of detecting non-speech boundaries between continuous speech segments (*col. 12, ln. 1-10, specifically elements 212 or 216 in figure 12*).

8. Regarding claim 28, Kimber et al. further disclose a method according to claim 25, wherein the step of detecting a speaker change is accomplished by use of at least one characteristic audio feature, in particular features derived from the spectrum of the audio signal (*col. 12, ln. 1-20, spectral feature vectors to train HMM are derived from audio signal for comparison with stored models*).

9. Regarding claim 29, Kimber et al. further disclose a method according to claim 25, wherein the step of performing a speaker recognition involves the particular steps of calculating a speaker signature from the audio stream and comparing the calculated speaker signature with at least one known speaker signature (*col. 5, ln. 10-27, spectral feature vectors used to train the HMM are speaker signatures*).

10. Regarding claim 30, Kimber et al. further disclose a method according to claim 25 for use in a speech recognition or voice control system comprising at least two speaker-specific speaker models and/or dictionaries, wherein interchanging between the at least two speaker-specific dictionaries dependent on the detected speaker change and the corresponding recognized speaker (*col. 11, ln. 13 to col. 12, ln. 20 and figure 9*).

11. Regarding claim 41, Kimber et al. further disclose an apparatus according to claim 39, further comprising a scanner which automatically scans a continuous audio record, in particular a continuous audio stream recorded on a data or a signal carrier, and for detecting speaker changes in the continuous audio record (*figure 11 or col. 11, ln. 13-37*).

12. Regarding claim 45, Kimber et al. further disclose an apparatus according to claim 39, comprising means for marking at least the beginning of a detected speech segment related to a predetermined speaker (*col. 11, ln. 21-37*).

13. Regarding claim 46, Kimber et al. further disclose an apparatus according to claim 39, comprising database, which stores speech signatures for at least two speakers (*the operation of figure 12 stores initial training speaker models*).

14. Regarding claim 47, Kimber et al. disclose a speech recognition processing an incoming audio stream containing human speech from a plurality of speakers and having at least two speaker models and/or speaker-specific dictionaries, comprising: a detector which detects a speaker change in the incoming audio stream (*col. 11, ln. 38-67*); a gather which gathers speaker-specific information with corresponding speaker-specific information of at least one predetermined known speaker from among the plurality of speakers thus recognizing the at least one predetermined speaker (*col. 5, ln. 10 to col. 10, ln. 67, input audio signal is parameterized into feature vectors for comparing with the speaker templates and/or col. 7, lines 41-67, recognizing speaker based on "usual speaking style of the speaker—the actual words used are unimportant"*); and an interchanger which interchanges between the at least two speaker-specific dictionaries dependent on the detected speaker change and the corresponding recognized speaker (*figure 12, the system of figure 12 contains a number of trained speaker recognized models, each is compared with input models to determine a speaker match*).

Claim Rejections - 35 USC § 103

15. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

16. Claims 19-24, 31-38, 42, and 48 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kimber et al. (US Patent No. 5598507) in view of Glickman et al. (US Patent No. 6076059).

17. Regarding claim 31, Kimber et al. disclose an apparatus for processing a continuous audio stream containing human speech from a plurality of speakers related to at least one particular transaction, comprising: a predeterminer which predetermines at least one known speaker from among the plurality of speakers (*element 212 in figure 12, initial training of speaker models HMMs*); a detector which detects speaker changes in the audio stream (*col. 11, ln. 38-67*); a recognizer which recognizes the predetermined speaker in the audio stream (*col. 12, ln.1- 27*).

Kimber et al. fail to disclose an initiator which initiates transcription of at least part of the audio stream in case of a detected speaker change and a recognized predetermined known speaker. However, Glickman et al. teach an initiator which initiates transcription of at least part of the audio stream in case of a detected speaker change and a recognized predetermined known speaker (*col. 5, ln. 30-67*).

Since Kimber et al. and Glickman et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Kimber et al. by incorporating the teaching of Glickman et al. in order to provide automatic closed-caption using speaker-dependent models to enhance speech recognition accuracy.

18. Regarding claims 19, 34, 42, and 48, Kimber et al. disclose a method, apparatus, and a program storage device readable by machine for processing a continuous audio stream containing human speech from a plurality of speakers related to at least one particular transaction, comprising the steps of: identifying a known speaker from among the plurality of speakers (*col. 7, lines 41-67, recognizing speaker based on "usual speaking style of the speaker—the actual words used are unimportant"*); digitizing the continuous audio stream (*ADC is inherently included in a computer system of figure 11*); detecting a speaker change in the digitized audio stream (*col. 11, ln. 38-67*); performing a speaker recognition if a speaker change is detected (*col. 12, ln.1- 27*).

Kimber et al. fail to disclose the step of transcribing at least part of the continuous audio stream if a predetermined speaker is recognized. However, Glickman et al. teach the step of transcribing at least part of the continuous audio stream if the known speaker is recognized (*col. 5, ln. 30-67*).

Since Kimber et al. and Glickman et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Kimber et al. by incorporating the teaching of Glickman et al. in order to provide automatic closed-caption using speaker-dependent models to enhance speech recognition accuracy.

19. Regarding claim 35, Kimber et al. disclose an apparatus for processing a continuous audio stream containing human speech related to at least one particular transaction, comprising the steps of: digitizing the continuous audio stream (*ADC is*

inherently included in a computer system of figure 11); detecting a speaker change in the digitized audio stream (*col. 11, ln. 38-67*); performing a speaker recognition if a speaker change is detected (*col. 11, ln. 38 to col. 12, ln. 27*); indexing the audio stream with respect to the detected speaker change if a predetermined speaker is recognized (*fig. 12 or col. 11, ln. 38 to col. 12, ln. 27*).

20. Regarding claims 20 and 36, Kimber et al. further disclose a method and apparatus according to claims 19 and 31, comprising the further step of protocolling time information for detected speaker changes (*col. 11, ln. 21-37*).

21. Regarding claims 21 and 32, Kimber et al. further disclose a method and apparatus according to claims 19 and 39, wherein the step of detecting a speaker change and/or the step of performing a speaker recognition is/are preceded by the further step of detecting non-speech boundaries between continuous speech segments (*col. 12, ln. 1-10, specifically elements 212 or 216 in figure 12*).

22. Regarding claim 22, Kimber et al. further disclose a method according to claim 19, wherein the step of detecting a speaker change is accomplished by use of at least one characteristic audio feature, in particular features derived from the spectrum of the audio signal (*col. 12, ln. 1-20, spectral feature vectors to train HMM are derived from audio signal for comparison with stored models*).

23. Regarding claim 23, Kimber et al. further disclose a method according to claim19, wherein the step of performing a speaker recognition involves the particular steps of calculating a speaker signature from the audio stream and comparing the calculated speaker signature with at least one known speaker signature (*col. 5, ln. 10-27, spectral feature vectors used to train the HMM are speaker signatures*).

24. Regarding claim 24, Kimber et al. further disclose a method according to claim 19 for use in a speech recognition or voice control system comprising at least two speaker-specific speaker models and/or dictionaries, wherein interchanging between the at least two speaker-specific dictionaries dependent on the detected speaker change and the corresponding recognized speaker (*col. 11, ln. 13 to col. 12, ln. 20 and figure 9*).

25. Regarding claim 33, Kimber et al. further disclose an apparatus according to claim 31, further comprising a scanner which automatically scans a continuous audio record, in particular a continuous audio stream recorded on a data or a signal carrier, and for detecting speaker changes in the continuous audio record (*figure 11 or col. 11, ln. 13-37*).

26. Regarding claim 37, Kimber et al. further disclose an apparatus according to claim 31, comprising means for marking at least the beginning of a detected speech segment related to a predetermined speaker (*col. 11, ln. 21-37*).

27. Regarding claim 38, Kimber et al. further disclose an apparatus according to claim 31, comprising database, which stores speech signatures for at least two speakers (*the operation of figure 12 stores initial training speaker models*).

Conclusion

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Kuhn et al. (US 6141644) is considered pertinent to the claimed invention.

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Huyen Vo whose telephone number is 703-305-8665. The examiner can normally be reached on M-F, 9-5:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached on 703-305-4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

HXV 5/23/2005

***

SUSAN MCFADDEN
PRIMARY EXAMINER